

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-180020
(43)Date of publication of application : 12.07.1996

(51)Int.CI. G06F 15/00
G06F 11/20

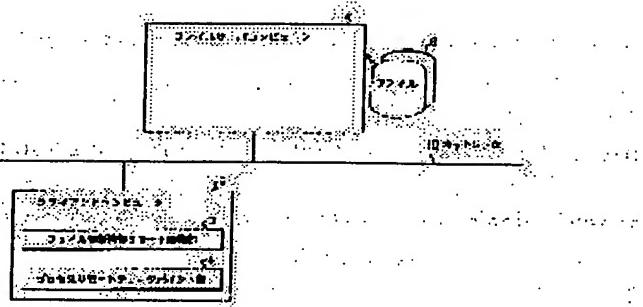
(21)Application number : 06-318494 (71)Applicant : TOSHIBA CORP
(22)Date of filing : 21.12.1994 (72)Inventor : HIRAYAMA HIDEAKI

(54) DISTRIBUTED PROCESSING SYSTEM

(57)Abstract:

PURPOSE: To improve the availability and reliability of a distributed processing system by re-executing a process under execution when the abnormality of a client computer is ended.

CONSTITUTION: When the client computer 2 which generated a fault is restarted, a file server computer 1 transmits file 9 updating information, the memory information of the process and the register information of a processor from the client computer 2 transmitted at a prescribed timing to the client computer 2. The client computer 2 which receives the transmission restarts the execution of the process which was during the execution at the time of fault generation from a check point indicated by the file 9 updating information, the memory information of the process and the register information of the processor. Thus, the availability and reliability of this distributed processing system are improved.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2000 Japan Patent Office

特開平8-180020

(43) 公開日 平成8年(1996)7月12日

(51) Int. C1.

G 06 F 15/00
11/20

識別記号 庁内整理番号

3 2 0 D 9364-5 L
3 1 0 A

F I

技術表示箇所

審査請求 未請求 請求項の数 4

○ L

(全 11 頁)

(21) 出願番号 特願平6-318494

(22) 出願日 平成6年(1994)12月21日

(71) 出願人 000003078

株式会社東芝

神奈川県川崎市幸区堀川町72番地

(72) 発明者 平山 秀昭

東京都青梅市末広町2丁目9番地 株式会社

東芝青梅工場内

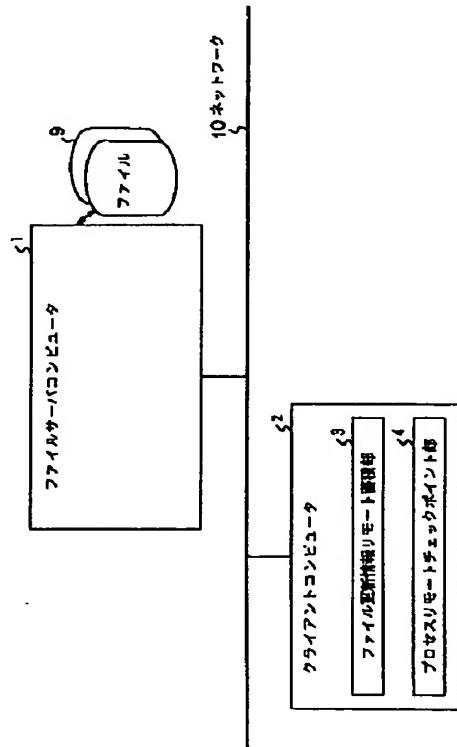
(74) 代理人 弁理士 鈴江 武彦

(54) 【発明の名称】分散処理システム

(57) 【要約】

【目的】異常終了した時に実行中であった処理を破棄せずに再実行することにより分散処理システムの信頼性をより向上させる分散処理装置を提供する。

【構成】ファイルサーバ機能を提供するファイルサーバコンピュータ1と、このファイルサーバコンピュータ1からファイルサーバ機能の提供を受けてプロセスを実行するクライアントコンピュータ2とから構成される分散処理システムにおいて、クライアントコンピュータ2に、ファイルの更新情報を蓄積するファイル更新情報リモート蓄積部3と、このファイル更新情報をプロセスのメモリ情報及びプロセッサのレジスタ情報とともに所定のタイミングでファイルサーバコンピュータ1に送信するプロセスリモートチェックポイント部4とを具備したことを特徴とする。



【特許請求の範囲】

【請求項1】 ファイルサーバ機能を提供するファイルサーバコンピュータと、このファイルサーバコンピュータからファイルサーバ機能の提供を受けてプロセスを実行するクライアントコンピュータとから構成される分散処理システムにおいて、

上記クライアントコンピュータに、ファイルの更新情報を蓄積するファイル更新情報リモート蓄積手段と、このファイル更新情報をプロセスのメモリ情報及びプロセッサのレジスタ情報をとともに所定のタイミングで上記ファイルサーバコンピュータに送信するプロセスリモートチェックポイント手段とを具備したことを特徴とする分散処理システム。

【請求項2】 ファイルサーバ機能を提供するファイルサーバコンピュータと、このファイルサーバコンピュータからファイルサーバ機能の提供を受けてプロセスを実行するクライアントコンピュータとから構成される分散処理システムにおいて、

上記ファイルサーバコンピュータに、上記クライアントコンピュータによるファイルの更新情報を蓄積するファイル更新情報ローカル蓄積手段と、上記クライアントコンピュータから所定のタイミングで送信されてくるプロセスのメモリ情報及びプロセッサのレジスタ情報を受信し蓄積するとともに、これらの情報と上記ファイル更新情報との同期を取得するプロセスローカルチェックポイント手段とを具備したことを特徴とする分散処理システム。

【請求項3】 上記クライアントコンピュータに、異常終了後、再立ち上げした際に、上記ファイルサーバコンピュータに蓄積されたファイル更新情報とプロセスのメモリ情報及びプロセッサのレジスタ情報を受信して、これらの情報が示すチェックポイントから処理を再開するプロセス再スタート手段を具備したことを特徴とする請求項1又は2記載の分散処理システム。

【請求項4】 上記クライアントコンピュータに、他のクライアントコンピュータが異常終了した際に、上記ファイルサーバコンピュータに蓄積された他のクライアントコンピュータのファイル更新情報とプロセスのメモリ情報及びプロセッサのレジスタ情報を受信して、これらの情報が示すチェックポイントから代替して処理を再開するプロセスマイグレーション手段を具備したことを特徴とする請求項1又は2記載の分散処理システム。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、例えばクライアントサーバシステム等の分散処理システムに適用して好適な分散処理装置に係り、特にクライアントコンピュータが異常終了したときに実行中であった処理を破棄せずに再実行することにより分散処理システムの信頼性をより向上させることを可能とする分散処理システムに関する。

【0002】

【従来の技術】 集中処理システムにおいて、集中処理を行っているコンピュータがシステムダウンを起こしてしまうと、集中処理システム全体が停止してしまうため、システムの可用性を向上させることができない。

【0003】 一方、従来の分散処理システムにおいては、システム内的一部のコンピュータが故障した場合であっても、他のコンピュータは正常に動作し続けることができるため、システム全体の可用性を向上させることが可能である。

【0004】 しかしながら、この場合においても、システムダウンを発生させてしまったコンピュータ上で実行中だったプロセスは、その際にアボートされてしまうために、システム全体の信頼性を向上させることができなかつた。

【0005】

【発明が解決しようとする課題】 上述したように、従来の分散処理システムにおいては、システムダウンを発生させてしまったコンピュータ上で実行中だったプロセスは、その際にアボートされてしまうために、システム全体の信頼性を向上させることができないという問題があった。

【0006】 本発明は上記実情に鑑みなされたものであり、異常終了した時に実行中であった処理を破棄せずに再実行することにより分散処理システムの信頼性をより向上させることを可能とする分散処理装置を提供することを目的とする。

【0007】

【課題を解決するための手段】 本発明は、ファイルサーバ機能を提供するファイルサーバコンピュータと、このファイルサーバコンピュータからファイルサーバ機能の提供を受けてプロセスを実行するクライアントコンピュータとから構成される分散処理システムにおいて、上記クライアントコンピュータに、ファイルの更新情報を蓄積するファイル更新情報リモート蓄積手段と、このファイル更新情報をプロセスのメモリ情報及びプロセッサのレジスタ情報をとともに所定のタイミングで上記ファイルサーバコンピュータに送信するプロセスリモートチェックポイント手段とを具備したことを特徴とする。

【0008】 また、本発明は、ファイルサーバ機能を提供するファイルサーバコンピュータと、このファイルサーバコンピュータからファイルサーバ機能の提供を受けてプロセスを実行するクライアントコンピュータとから構成される分散処理システムにおいて、上記ファイルサーバコンピュータに、上記クライアントコンピュータによるファイルの更新情報を蓄積するファイル更新情報ローカル蓄積手段と、上記クライアントコンピュータから所定のタイミングで送信されてくるプロセスのメモリ情報及びプロセッサのレジスタ情報を受信し蓄積するとともに、これらの情報と上記ファイル更新情報との同期を

取得するプロセスローカルチェックポイント手段とを具備したことを特徴とする。

【0009】また、本発明は、上記クライアントコンピュータに、異常終了後、再立ち上げした際に、上記ファイルサーバコンピュータに蓄積されたファイル更新情報とプロセスのメモリ情報及びプロセッサのレジスタ情報とを受信して、これらの情報が示すチェックポイントから処理を再開するプロセス再スタート手段を具備したことを特徴とする。

【0010】また、本発明は、上記クライアントコンピュータに、他のクライアントコンピュータが異常終了した際に、上記ファイルサーバコンピュータに蓄積された他のクライアントコンピュータのファイル更新情報とプロセスのメモリ情報及びプロセッサのレジスタ情報を受信して、これらの情報が示すチェックポイントから代替して処理を再開するプロセスマイグレーション手段を具備したことを特徴とする。

【0011】

【作用】本発明の構成によれば、クライアントコンピュータは、常時、ファイル更新情報リモート蓄積手段によりファイルの更新情報を蓄積する。そして、プロセスリモートチェックポイント手段が、このファイル更新情報とプロセスのメモリ情報及びプロセッサのレジスタ情報を所定のタイミングでファイルサーバコンピュータに送信する。

【0012】ここで、クライアントコンピュータに障害が発生した場合を考えると、まず、この障害を発生させたクライアントコンピュータが再立ち上げされた際に、ファイルサーバコンピュータが、上記所定のタイミングで送信された当該クライアントコンピュータからのファイル更新情報と、プロセスのメモリ情報及びプロセッサのレジスタ情報をこのクライアントコンピュータに送信する。

【0013】そして、この送信を受信したクライアントコンピュータは、このファイル更新情報とプロセスのメモリ情報及びプロセッサのレジスタ情報とが示すチェックポイントから障害発生時に実行していたプロセスの実行を再開する。

【0014】これにより、障害発生時に実行していたプロセスの処理をアボートさせることなく再開することができるため、分散処理システムの信頼性をより向上させることができる。

【0015】また、システムの設定によっては、障害を発生させたクライアントコンピュータの再立ち上げを待機せずに、例えばホットスタンバイ等の他のクライアントコンピュータに、障害を発生させたクライアントコンピュータのファイル更新情報とプロセスのメモリ情報及びプロセッサのレジスタ情報を送信し、この他のクライアントコンピュータに継続して代替実行させる。

【0016】これにより、上述と同様に障害発生時に実

行していたプロセスの処理をアボートさせることなく再開することができるため、分散処理システムの信頼性をより向上させることができるとなる。

【0017】また、本発明の構成によれば、ファイルサーバコンピュータは、ファイル更新情報ローカル蓄積手段によりクライアントコンピュータにより行われたファイルの更新情報を蓄積していく。

【0018】また、プロセスローカルチェックポイント手段が、クライアントコンピュータから所定のタイミングで送信されてくるプロセスのメモリ情報やプロセッサのレジスタ情報を受信し蓄積するとともに、これらの情報とファイル更新情報との同期を取得する。

【0019】ここで、クライアントコンピュータに障害が発生した場合を考えると、まず、この障害を発生させたクライアントコンピュータが再立ち上げされた際に、ファイルサーバコンピュータが、上記所定のタイミングで送信された当該クライアントコンピュータからのプロセスのメモリ情報及びプロセッサのレジスタ情報と、これらの情報と同期したファイル更新情報をこのクライアントコンピュータに送信する。

【0020】そして、この送信を受信したクライアントコンピュータは、このファイル更新情報とプロセスのメモリ情報及びプロセッサのレジスタ情報とが示すチェックポイントから障害発生時に実行していたプロセスの実行を再開する。

【0021】これにより、障害発生時に実行していたプロセスの処理をアボートさせることなく再開することができるため、分散処理システムの信頼性をより向上させることができる。

【0022】また、システムの設定によっては、障害を発生させたクライアントコンピュータの再立ち上げを待機せずに、例えばホットスタンバイ等の他のクライアントコンピュータに、障害を発生させたクライアントコンピュータのプロセスのメモリ情報及びプロセッサのレジスタ情報を、これらの情報と同期したファイル更新情報を送信し、この他のクライアントコンピュータに継続して代替実行させる。

【0023】これにより、上述と同様に障害発生時に実行していたプロセスの実行をアボートさせることなく再開することができるため、分散処理システムの信頼性をより向上させることができる。

【0024】

【実施例】以下図面を参照して本発明の一実施例を説明する。まず、図1を参照して本発明の第1実施例に係る分散処理システムの概略構成を説明する。

【0025】図1は本発明の第1実施例に係る分散処理システムの概略構成図である。同実施例に係る分散処理システムは、図1に示すように、ファイルサーバコンピュータ1とクライアントコンピュータ2とがネットワーク10を介して接続された構成となっている。この分散

処理システムで使用されるファイル9は、ファイルサーバコンピュータ1により管理されており、また、通常のプロセスは、クライアントコンピュータ2上で実行されている。そして、このプロセスによるファイルアクセスは、ファイルサーバコンピュータ1によるファイルサーバ機能によって実現されている。

【0026】ただし、クライアントコンピュータ2上のプロセスによるファイルの更新は、ファイル更新情報リモート蓄積部3によりその更新情報が蓄積されるのみであり、コンピュータ1が管理する実際のファイル9への更新は行われない。そして、この実際のファイルの更新は、所定のタイミングで、プロセスリモートチェックポイント部4からのファイル更新情報の送信により行う。また、この時、同時にプロセスのメモリ情報やプロセッサのレジスタ情報をファイルサーバコンピュータ1上に送信することによりチェックポイントを取得する。

【0027】次に、図2を参照して本発明の第2実施例に係る分散処理システムの概略構成を説明する。図2は本発明の第2実施例に係る分散処理システムの概略構成図である。

【0028】同実施例に係る分散処理システムは、図2に示すように、クライアントコンピュータ2によるファイル9の更新について、その更新情報がファイルサーバコンピュータ1のファイル更新情報ローカル蓄積部5によって蓄積される。また、実際のファイルの更新は、所定のタイミングでクライアントコンピュータ2からプロセスのメモリ情報やプロセッサのレジスタ情報を送信されてきた際に、プロセスローカルチェックポイント部6が、これら的情報の受信と同期をとつて行う。即ち、これによりクライアントコンピュータ2のチェックポイントが取得されることとなる。

【0029】ここで、図3及び図4を参照して第1及び第2実施例の動作原理を説明する。図3及び図4は、第1及び第2実施例の動作原理を説明するための概念図である。

【0030】上述したように、第1及び第2実施例のいずれかの構成によれば、プロセスの実行途中の状態をチェックポイントとしてファイルサーバコンピュータ1上に保持しているために、仮にプロセスの実行途中でクライアントコンピュータ2がシステムダウンを発生させたとしても、図3に示すように、障害を発生させたクライアントコンピュータ2を再立ち上げした後で、プロセス再スタート部7にファイル更新情報とプロセスのメモリ情報及びプロセッサのレジスタ情報を送信することによって、障害発生時に実行中だったプロセスを再実行することが可能となる。

【0031】またファイルサーバコンピュータ1が異常終了した場合であっても、これらのチェックポイントがファイルサーバコンピュータ1の安定記憶装置上に保持されているのであれば、ファイルサーバコンピュータ1

の障害発生時にコンピュータ2上で実行中だったプロセスを一度アポートし、チェックポイントから再実行することにより正常に処理し続けることが可能となる。

【0032】またプロセスの実行途中の状態をチェックポイントとしてファイルサーバコンピュータ1上に保持しているので、システムの設定によっては、図4に示すように、もしプロセスの実行途中でプロセスを実行するコンピュータ2aが異常終了したとしても、例えばホットスタンバイ等の別のコンピュータ2bのプロセスマイグレーション部8に、コンピュータ2aについてのファイル更新情報と、プロセスのメモリ情報及びプロセッサのレジスタ情報を送信することにより、コンピュータ2aが異常終了したときに実行中だったプロセスをコンピュータ2b上で継続して代替実行せざるを得なくなる。

【0033】次に、図5乃至図11を参照して第1及び第2実施例の動作手順を説明する。まず、図5を参照して同実施例におけるプロセス処理の動作手順を説明する。図5は同実施例におけるプロセス処理の動作手順を説明するためのフローチャートである。

【0034】まず、このプロセスの処理が、ファイルの更新処理か否かを判断する(図5のステップA1)。ここで、ファイルの更新処理であった場合には(図5のステップA1のY)、該当のファイルをまず更新し(図5のステップA2)、ファイル更新情報リモート蓄積部3によってファイル更新情報をクライアントコンピュータ2に蓄積するか(図1に示す構成の場合)、ファイル更新情報ローカル蓄積部5によってファイル更新情報をファイルサーバコンピュータ1に蓄積する(図2に示す構成の場合)(図5のステップA3)。

【0035】一方、このプロセスの処理が、ファイルの更新処理でなかった場合には(図5のステップA1のN)、通常通りの処理を行う(図5のステップA4)。次に、図6を参照して同実施例におけるチェックポイント処理の動作手順を説明する。

【0036】図6は同実施例におけるチェックポイント処理の動作手順を説明するためのフローチャートである。まず、ファイル更新情報リモート蓄積部3によってクライアントコンピュータ2に蓄積されたファイル更新情報をファイルサーバコンピュータ1に送信することにより実際のファイルに反映するか(図1に示す構成の場合)、ファイル更新情報ローカル蓄積部5によってファイルサーバコンピュータ1に蓄積されたファイル更新情報を実際のファイルに反映する(図2に示す構成の場合)(図6のステップB1)。

【0037】そして、このタイミングと同期して、クライアントコンピュータ2のプロセスのメモリ情報及びプロセッサのレジスタ情報をファイルサーバコンピュータ1に送信して記録する(図6のステップB2)。

【0038】次に、図7を参照してクライアントコンピ

ュータに障害が発生した際、そのコンピュータの再立ち上げ後に処理を再開させる場合の動作手順を説明する。図7はクライアントコンピュータに障害が発生した際、そのコンピュータの再立ち上げ後に処理を再開させる場合の動作手順を説明するためのフローチャートである。

【0039】ここでは、同時に図3を参照して説明する。まず、障害を発生させたクライアントコンピュータ2を再スタートする(図7のステップC1)。次に、ファイルサーバコンピュータ1上に記録されている、障害発生時に実行中だったプロセスのファイル更新情報と、プロセスのメモリ情報及びプロセッサのレジスタ情報を再スタートしたクライアントコンピュータ2に送信する。そして、これらの情報からプロセス再スタート部7により再実行を実施する(図7のステップC2)。

【0040】次に、図8を参照してクライアントコンピュータに障害が発生した際に、他のコンピュータに継続して代替実行させる場合の動作手順を説明する。図8はクライアントコンピュータに障害が発生した際に、他のコンピュータに継続して代替実行させる場合の動作手順を説明するためのフローチャートである。

【0041】ここでは、同時に図4を参照して説明する。この場合、ファイルサーバコンピュータ1は、障害を発生させたクライアントコンピュータ2aとは別のクライアントコンピュータ2bに、障害を発生させたクライアントコンピュータ2aのファイル更新情報と、プロセスのメモリ情報及びプロセッサのレジスタ情報を送信する。そして、障害発生時に実行中だったプロセスをクライアントコンピュータ2bのプロセスマイグレーション部8に代替実行することにより再実行する(図8のステップD1)。

【0042】次に、図9を参照してファイルサーバコンピュータに障害が発生した場合の動作手順を説明する。図9はファイルサーバコンピュータに障害が発生した場合の動作手順を説明するためのフローチャートである。

【0043】ここでは、同時に図3を参照して説明する。まず、ファイルサーバコンピュータ1を再スタートし(図9のステップE1)、このときクライアントコンピュータ2で実行中のプロセスをアポートする(図9のステップE2)。

【0044】そして、このアポートしたプロセスを、プロセス再スタート部7によりファイルサーバコンピュータの安定記録装置上に記録されているチェックポイントから再実行する(図9のステップE3)。

【0045】次に、図10を参照してプロセスの実行を継続したままオペレーティングシステム(以下OSという)のバージョンを更新する第1の動作手順を説明する。図10はプロセスの実行を継続したままOSのバージョンを更新する第1の動作手順を説明するためのフローチャートである。

【0046】ここでは、同時に図3を参照して説明す

る。まず、クライアントコンピュータ2上で実行中のプロセスのチェックポイントを取得し(図10のステップF1)、これらのプロセスを全てアポートする(図10のステップF2)。そして、このクライアントコンピュータ2を一度シャットダウンする(図10のステップF3)。

【0047】ここで、クライアントコンピュータ2を再立ち上げする際に、新バージョンのOSで再スタートを実施する(図10のステップF4)。そして、クライアントコンピュータ2でアポート時に実行中だったプロセスを、プロセス再スタート部7によりファイルサーバコンピュータ1上に記録されているチェックポイントから再実行する(図10のステップF4)。

【0048】これにより、クライアントコンピュータは、プロセスの処理を継続したままOSのバージョンを新しくすることが可能となる。次に、図11を参照してプロセスの実行を継続したままOSのバージョンを更新する第2の動作手順を説明する。

【0049】図11はプロセスの実行を継続したままOSのバージョンを更新する第2の動作手順を説明するためのフローチャートである。ここでは、同時に図4を参照して説明する。

【0050】まず、クライアントコンピュータ2a上で実行中のプロセスを、別のクライアントコンピュータ2bのマイグレーション部8により代替実行させる(図11のステップG1)。これは、ファイルサーバコンピュータ1に保持するクライアントコンピュータ2aのチェックポイントを、クライアントコンピュータ2bに引き渡すことにより行う。

【0051】次に、クライアントコンピュータ2aを一度シャットダウンする(図11のステップG2)。ここで、クライアントコンピュータ2aを再立ち上げする際に、新バージョンのOSで再スタートを実施する(図11のステップG3)。そして、別のクライアントコンピュータ2bに代替実行させたプロセスを、新バージョンのOSで再スタートしたコンピュータに再度戻す。

【0052】これにより、第1の処理と同様、クライアントコンピュータは、プロセスの処理を継続したままOSのバージョンを新しくすることが可能となる。なお、ファイルサーバコンピュータ1がフォールトトレランコンピュータで構成された場合、システム全体をフォールトトレランとすることも可能である。

【0053】

【発明の効果】以上詳記したように、本発明によれば、異常終了した時に実行中であった処理を破棄せずに再実行することにより、分散処理システムの可用性と共に信頼性を向上させることができ、高信頼性、高可用性の分散処理システムを実現することが可能となる。

【図面の簡単な説明】

【図1】本発明の第1実施例に係る分散処理システムの

概略構成図。

【図2】本発明の第2実施例に係る分散処理システムの概略構成図。

【図3】第1及び第2実施例の動作原理を説明するための概念図。

【図4】第1及び第2実施例の動作原理を説明するための概念図。

【図5】同実施例におけるプロセス処理の動作手順を説明するためのフローチャート。

【図6】同実施例におけるチェックポイント処理の動作手順を説明するためのフローチャート。

【図7】クライアントコンピュータに障害が発生した際、そのコンピュータの再立ち上げ後に処理を再開させる場合の動作手順を説明するためのフローチャート。

【図8】クライアントコンピュータに障害が発生した際に、他のコンピュータに継続して代替実行させる場合の動作手順を説明するためのフローチャート。

【図9】ファイルサーバコンピュータに障害が発生した場合の動作手順を説明するためのフローチャート。

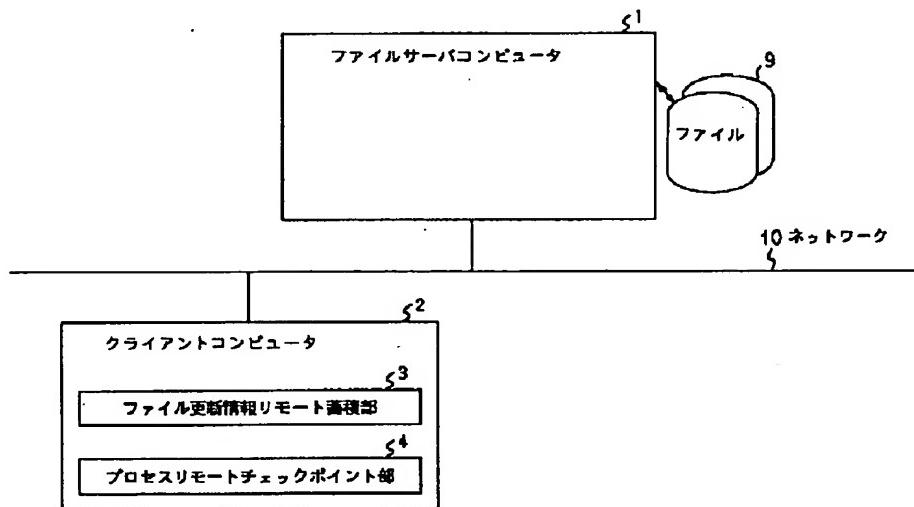
【図10】プロセスの実行を継続したままオペレーティングシステムのバージョンを更新する第1の動作手順を説明するためのフローチャート。

【図11】プロセスの実行を継続したままオペレーティングシステムのバージョンを更新する第2の動作手順を説明するためのフローチャート。

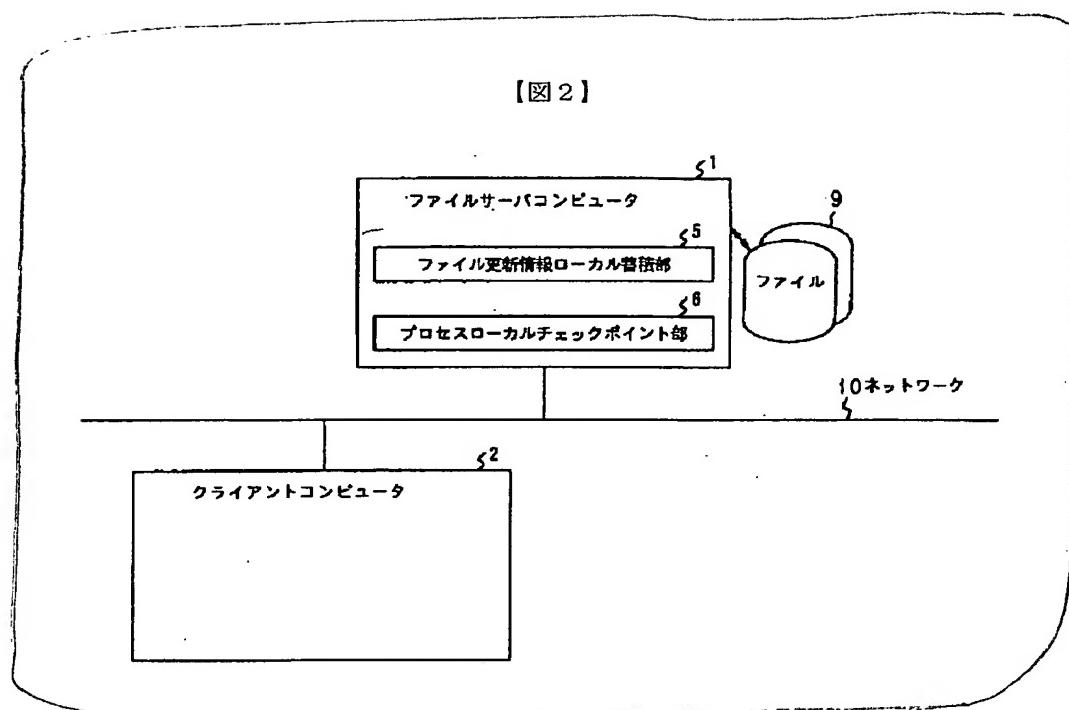
【符号の説明】

- 1…ファイルサーバコンピュータ、2, 2a, 2b…クライアントコンピュータ、3…ファイル更新情報リモート蓄積部、4…プロセスリモートチェックポイント部、5…ファイル更新情報ローカル蓄積部、6…プロセスローカルチェックポイント部、7…プロセス再スタート部、8…プロセスマイグレーション部、9…ファイル、10…ネットワーク。

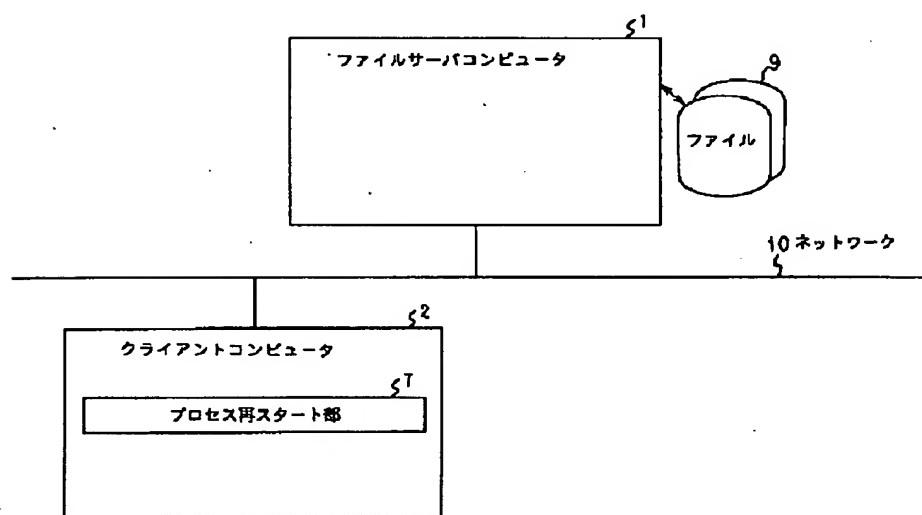
【図1】



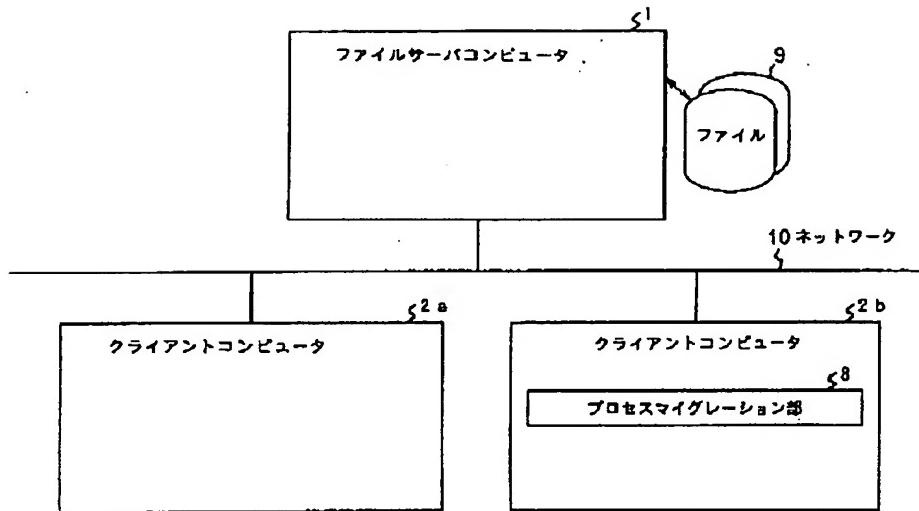
【図2】



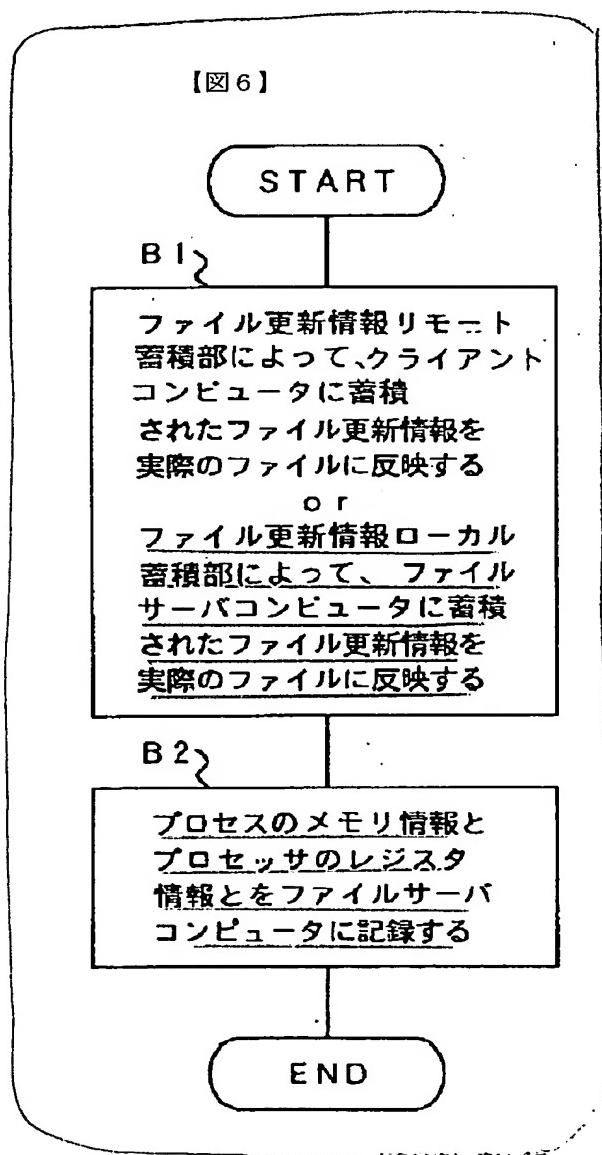
【図3】



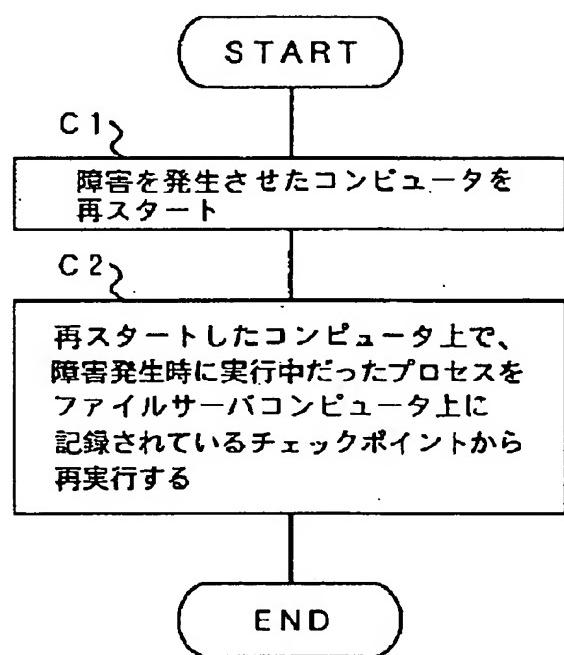
【図4】



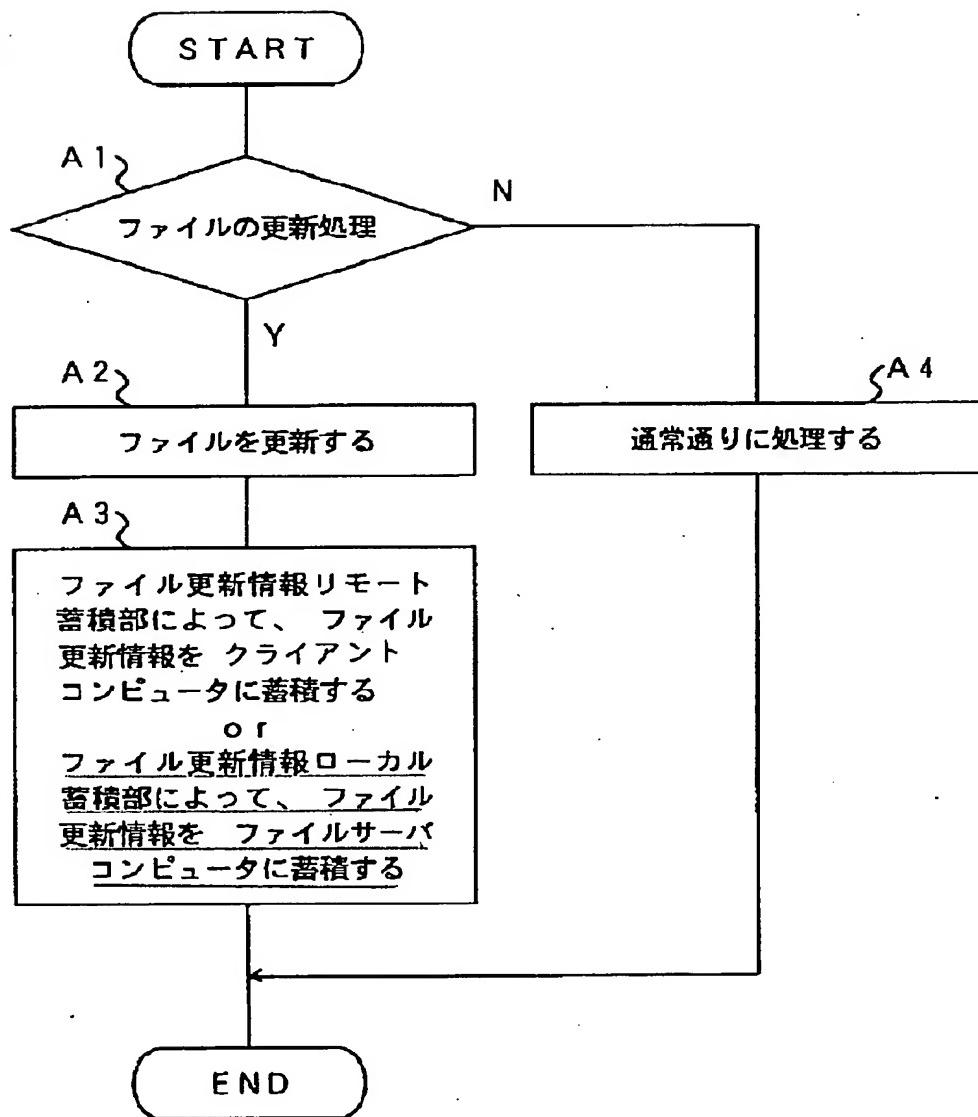
【図6】



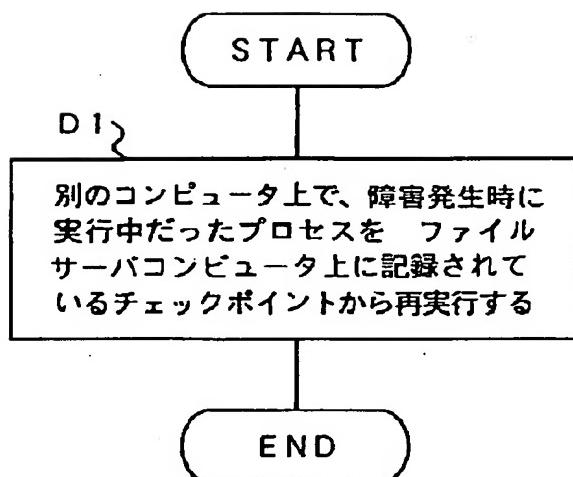
【図7】



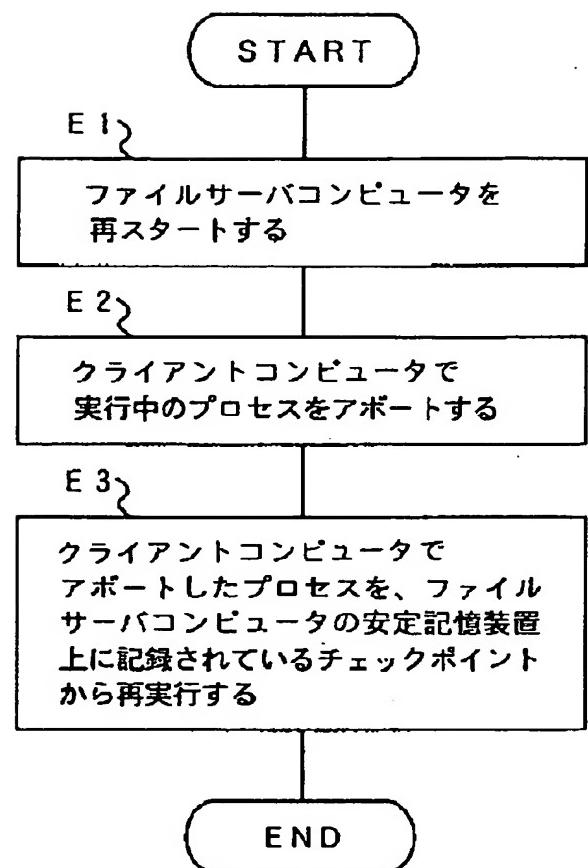
【図5】



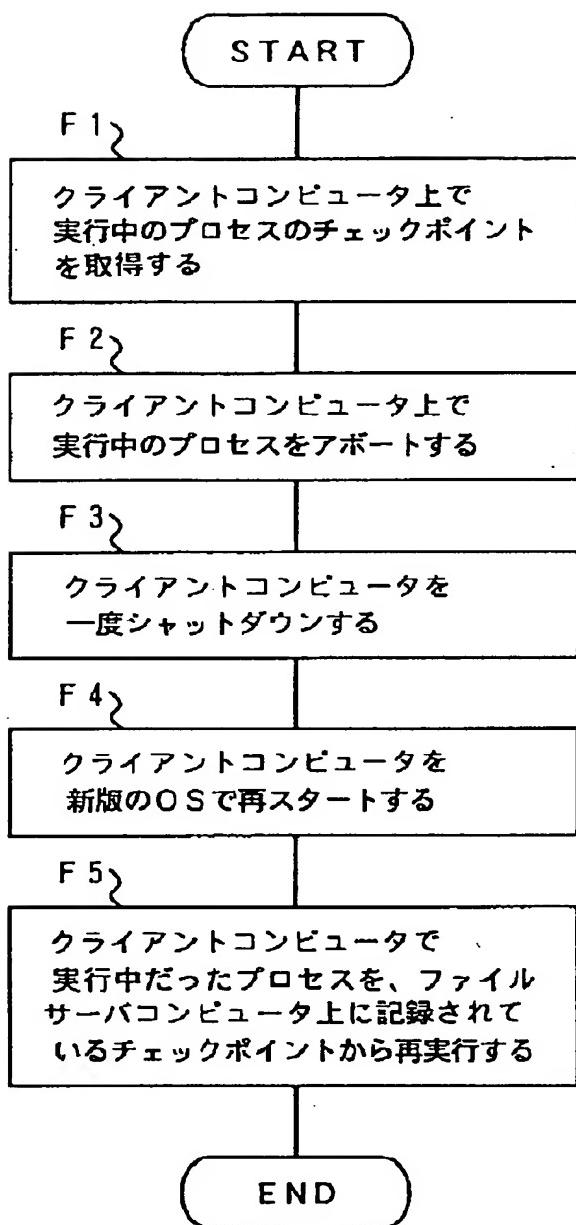
【図8】



【図9】



【図10】



【図11】

